Lecture 1: Sets, Functions, Logic, and Proofs

1 Sets

1.1 Definition and Operations

A *set* is a collection of objects which we call *elements*. We typically denote sets by capital letters and elements by lowercase letters. If x is an element of a set A, we write $x \in A$. If x is not an element of A, we write $x \notin A$. For example, if $A = \{1, 2, 3\}$, then $1 \in A$ and $4 \notin A$.

There are multiple ways to describe a set. One is to list its elements, as in $A = \{1, 2, 3\}$. Another is to describe a property that characterizes its elements, as in $B = \{x \in \mathbb{R} : x > 0\}$, which is the set of positive real numbers.

Some examples of sets include:

- The empty set $\emptyset$.
- The set of integers $\mathbb{Z} = \{\dots, -2, -1, 0, 1, 2, \dots\}$.
- The set of real numbers $\mathbb{R}$.
- The set of natural numbers $\mathbb{N} = \{1, 2, 3, \dots\}$.
- The set of rational numbers $\mathbb{Q} = \{p/q : p \in \mathbb{Z}, q \in \mathbb{Z} \setminus \{0\}\}$.
- The set of all subsets of a finite set S, called the power set and denoted 2^S.

Definition 1.1 (Inclusion and equality)**.** For sets A and B, write $A \subseteq B$ (A is a *subset* of B) if every element of A belongs to B. The sets are equal if $A \subseteq B$ and $B \subseteq A$. Write $A \subset B$ (A is a *proper subset* of B) if $A \subseteq B$ and $A \neq B$.

For example, the statements

$$\{1, 2\} \subseteq \{1, 2, 3\}, \qquad \{1, 2, 3\} = \{3, 2, 1\}, \qquad \mathbb{Z} \subset \mathbb{R}, \qquad \mathbb{N} \subset \mathbb{Z}$$

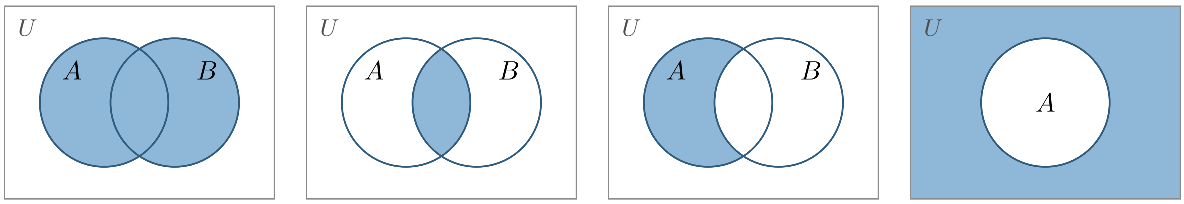
are all true. Furthermore, for any set S, we have $\emptyset \subseteq S$ and $S \subseteq S$.

Definition 1.2 (Set operations)**.** For sets A and B, the *union*, *intersection*, and *set difference* are

$$\begin{aligned} A \cup B &= \{x : x \in A \text{ or } x \in B\}, \\ A \cap B &= \{x : x \in A \text{ and } x \in B\}, \\ A \setminus B &= \{x \in A : x \notin B\}. \end{aligned}$$

The *complement* of a set A relative to a universal set U is $A^c = \{x : x \in U \text{ and } x \notin A\} = U \setminus A$.

These operations are called union, intersection, and set difference. They are typically illustrated by Venn diagrams:



$A \cup B$ $\qquad$ $A \cap B$ $\qquad$ $A \setminus B$ $\qquad$ $A^c = U \setminus A$

Figure 1.1: Union, intersection, set difference, and complement.

Of course, we can also define the union and intersection of more than two sets. Given a collection of sets $\{A_i\}_{i \in I}$ indexed by a set I, we define

$$\bigcup_{i \in I} A_i = \{x : x \in A_i \text{ for some } i \in I\}, \qquad \bigcap_{i \in I} A_i = \{x : x \in A_i \text{ for every } i \in I\}.$$

The following theorem summarizes some properties of aforementioned set operations. Its proof is left as an exercise.

Theorem 1.3 (Set algebra)**.** *For sets A, B, and C,*

1. $A \cup B = B \cup A$, $A \cap B = B \cap A$ *(commutativity),*
2. $A \cup (B \cup C) = (A \cup B) \cup C$, $A \cap (B \cap C) = (A \cap B) \cap C$ *(associativity),*
3. $A \cup (A \cap B) = A$, $A \cap (A \cup B) = A$ *(absorption),*
4. $A \setminus B = A \cap B^c$ *(set difference),*
5. $(A \cup B)^c = A^c \cap B^c$, $(A \cap B)^c = A^c \cup B^c$ *(De Morgan's laws).*

We say that sets are *disjoint* if their intersection is empty. For example, the sets of even and odd integers are disjoint. A *partition* of a set S is a collection of nonempty, pairwise disjoint subsets of S whose union is S. For example, the sets of even and odd integers form a partition of $\mathbb{Z}$.

1.2 Cartesian Product and $\mathbb{R}^n$

Definition 1.4. A *Cartesian product* of two sets A and B is the set

$$A \times B = \{(a, b) : a \in A, b \in B\}.$$

For example, if $A = \{1, 2\}$ and $B = \{x, y\}$, then

$$A \times B = \{(1, x), (1, y), (2, x), (2, y)\}.$$

The product $A_1 \times \cdots \times A_n$ consists of tuples $(a_1, \ldots, a_n)$ satisfying $a_i \in A_i$ for every i. An n-fold product of a set A with itself is denoted $A^n = A \times \cdots \times A$.

The most commonly used in economics example of an n-fold product is $\mathbb{R}^n$. In this course, we treat elements of $\mathbb{R}^n$ as column vectors:

$$x = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} \in \mathbb{R}^n,$$

but that convention is not important until we discuss linear algebra.

Sometimes we need to compare vectors in $\mathbb{R}^n$. While not all vectors are comparable, we can compare them coordinatewise.

Definition 1.5 (Coordinatewise comparisons)**.** For $x, y \in \mathbb{R}^n$, write

$$\begin{aligned} x \geq y &\text{ if } x_i \geq y_i \text{ for every } i, \\ x > y &\text{ if } x \geq y \text{ and } x \neq y, \\ x \gg y &\text{ if } x_i > y_i \text{ for every } i. \end{aligned}$$

The nonnegative and positive orthants are

$$\mathbb{R}^n_+ = \{x \in \mathbb{R}^n : x \geq 0\}, \qquad \mathbb{R}^n_{++} = \{x \in \mathbb{R}^n : x \gg 0\}.$$

2 Functions

2.1 Vocabulary and properties

Definition 2.1 (Function)**.** Let X and Y be two sets. A *function* $f : X \to Y$ is a rule that assigns to each $x \in X$ exactly one element $f(x) \in Y$. We also call f a *map*, or a *mapping*, from X to Y.

For a function $f : X \to Y$, we call X the *domain* and Y the *codomain*. For a subset $A \subseteq X$ of the domain, the *image* of A is the set

$$f(A) = \{f(x) : x \in A\} \subseteq Y$$

of values taken by f on A. In particular, we call $f(X)$ the *range* of f. The *preimage* of a subset $B \subseteq Y$ of the codomain is the set

$$f^{-1}(B) = \{x \in X : f(x) \in B\} \subseteq X$$

of points in the domain that map into B. The *graph* of f is the set of pairs

$$\{(x, f(x)) : x \in X\} \subseteq X \times Y.$$

Definition 2.2 (Injective, surjective, bijective)**.** A function $f : X \to Y$ is

1. *injective*, or *one-to-one*, if $f(x) = f(x')$ implies $x = x'$ for all $x, x' \in X$;
2. *surjective*, or *onto*, if its range is all of Y, that is, $f(X) = Y$;
3. *bijective* if it is both injective and surjective.

Injectivity says that distinct points of the domain have distinct values, and surjectivity says that every $y \in Y$ equals $f(x)$ for at least one $x \in X$. The two conditions are checked in opposite directions: injectivity starts from two points of X and compares their values, while surjectivity starts from a point of Y and asks for a point of X that reaches it.

Note that whether a function is surjective depends on the declared domain and codomain. For example, the function $f(x) = x^2$ is not surjective if $f : \mathbb{R} \to \mathbb{R}$ but surjective if $f : \mathbb{R} \to \mathbb{R}_+$. On the other hand, the function $g(x) = x^3$ is surjective if $g : \mathbb{R} \to \mathbb{R}$ or $g : \mathbb{R}_+ \to \mathbb{R}_+$.

Definition 2.3 (Composition)**.** Let $f : X \to Y$ and $g : Y \to Z$, so that the codomain of f is the domain of g. The *composition* of f and g is the function $g \circ f : X \to Z$ defined by

$$(g \circ f)(x) = g(f(x)) \quad \text{for every } x \in X.$$

2.2 Inverse Function

Let $f : X \to Y$ be a function and let $y \in Y$. Recalling the definition of a preimage, let us consider the set $f^{-1}(\{y\})$, which contains all elements of X that map to the point y. In general, this set may be empty, contain one point, or contain multiple points. By Definition 2.2, if f is surjective, then $f^{-1}(\{y\})$ contains at least one point for every $y \in Y$, and if f is injective, then it contains at most one point. Consequently, if f is bijective, then it contains exactly one point. This allows us to slightly abuse notation and define $f^{-1}(y)$ to be the unique point of X that maps to y, so that $f^{-1}(\{y\}) = \{f^{-1}(y)\}$.

Definition 2.4 (Inverse)**.** Let $f : X \to Y$ be a bijective function. The *inverse* of f is the function $f^{-1} : Y \to X$ that assigns to each $y \in Y$ the unique $x \in X$ with $f(x) = y$.

The function f and its inverse f^{-1} undo each other, in the sense that

$$f^{-1}(f(x)) = x \text{ for every } x \in X, \qquad f(f^{-1}(y)) = y \text{ for every } y \in Y.$$

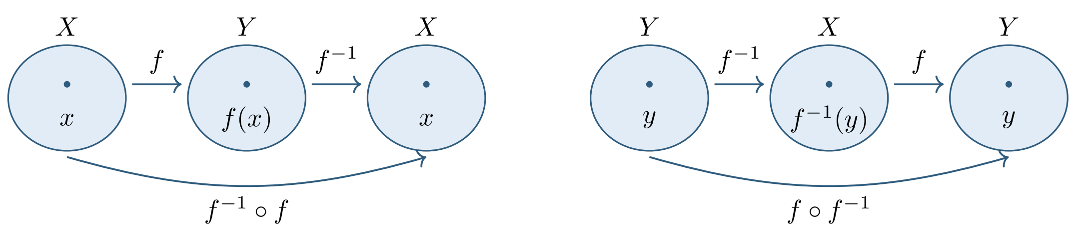


Figure 2.1: Applying f and then f^{-1} returns every point of X to itself; applying f^{-1} and then f returns every point of Y to itself.

The theorem below shows that if we tried to define an "inverse" for a function that is not bijective, we would fail.

Theorem 2.5 (Existence of an inverse)**.** *For a function $f : X \to Y$, the following are equivalent.*

1. *f is bijective.*
2. *There is a function $g : Y \to X$ with $(g \circ f)(x) = x$ for every $x \in X$ and $(f \circ g)(y) = y$ for every $y \in Y$.*

Moreover, the function g in (2) is unique and equals f^{-1}.

Proof (optional). An equivalence asserts two implications, so we prove that (1) implies (2) and that (2) implies (1). The uniqueness claim is proved last.

(1) implies (2). Assume f is bijective. Then $f^{-1} : Y \to X$ is defined by Definition 2.4, and we take $g = f^{-1}$. Let $x \in X$ and put $y = f(x)$. By Definition 2.4, $f^{-1}(y)$ is the unique point of X that f maps to y; since $f(x) = y$, that point is x. Using Definition 2.3 to expand the composition,

$$(f^{-1} \circ f)(x) = f^{-1}(f(x)) = f^{-1}(y) = x.$$

Now let $y \in Y$ and put $x = f^{-1}(y)$, so that $f(x) = y$ by Definition 2.4. Then

$$(f \circ f^{-1})(y) = f(f^{-1}(y)) = f(x) = y.$$

Both equations hold for arbitrary $x \in X$ and $y \in Y$, so $g = f^{-1}$ is a function of the kind required in (2).

(2) implies (1). Assume such a g exists. To show that f is injective, let $x, x' \in X$ satisfy $f(x) = f(x')$. Applying g to both sides and using Definition 2.3,

$$x = (g \circ f)(x) = g(f(x)) = g(f(x')) = (g \circ f)(x') = x',$$

where the first and last equalities are the hypothesis on g. So f is injective in the sense of Definition 2.2. To show that f is surjective, let $y \in Y$ and put $x = g(y)$, which lies in X because $g : Y \to X$. Then

$$f(x) = f(g(y)) = (f \circ g)(y) = y,$$

so $y \in f(X)$. Since y was an arbitrary point of Y, this gives $Y \subseteq f(X)$, and the reverse inclusion $f(X) \subseteq Y$ holds because f takes its values in Y. Hence $f(X) = Y$, which is surjectivity in Definition 2.2. Being both injective and surjective, f is bijective.

Uniqueness. Let g be any function satisfying (2). By the previous paragraph f is bijective, so f^{-1} is defined. Let $y \in Y$ and put $x = f^{-1}(y)$, so that $f(x) = y$ by Definition 2.4. Then

$$g(y) = g(f(x)) = (g \circ f)(x) = x = f^{-1}(y).$$

The functions g and f^{-1} have the same domain Y and the same codomain X, and we have just shown that they agree at every $y \in Y$. Hence $g = f^{-1}$. □

3 Logic

3.1 Logical Statements and Operations

A *statement* (also referred to as a *proposition* or an *assertion*) is a sentence that is either true or false. For example, "$2 + 2 = 4$" is a true statement and "$2 + 2 = 5$" is a false one.

Given a statement P, its *negation* $\neg P$ is true if and only if P is false. Given two statements P and Q, their *conjunction* $P \wedge Q$ is true if and only if both P and Q are true, while their *disjunction* $P \vee Q$ is true if and only if at least one of P or Q is true. We typically illustrate these logical operations with truth tables:

P	$\neg P$
T	F
F	T

P	Q	$P \wedge Q$
T	T	T
T	F	F
F	T	F
F	F	F

P	Q	$P \vee Q$
T	T	T
T	F	T
F	T	T
F	F	F

3.2 Implication and equivalence

Perhaps the most important logical operation is implication. We define it via a truth table as follows:

P	Q	$P \Rightarrow Q$
T	T	T
T	F	F
F	T	T
F	F	T

The implication $P \Rightarrow Q$ is read "P implies Q," or "if P, then Q," or "P is sufficient for Q," or "Q is necessary for P." Its *converse* is $Q \Rightarrow P$, and its *contrapositive* is $\neg Q \Rightarrow \neg P$. If $P \Rightarrow Q$ and $Q \Rightarrow P$ both hold, we write $P \Leftrightarrow Q$ and say that P and Q are *equivalent*, or P if and only if (IFF) Q.

Our first proof involves showing that an implication and its contrapositive are equivalent.

Theorem 3.1 (Contrapositive law)**.**

$$(P \Rightarrow Q) \iff (\neg Q \Rightarrow \neg P).$$

Proof. The truth tables for $P \Rightarrow Q$ and $\neg Q \Rightarrow \neg P$ are

P	Q	$P \Rightarrow Q$
T	T	T
T	F	F
F	T	T
F	F	T

P	Q	$\neg Q$	$\neg P$	$\neg Q \Rightarrow \neg P$
T	T	F	F	T
T	F	T	F	F
F	T	F	T	T
F	F	T	T	T

Since the truth tables are the same, the statements are equivalent. □

Note that an implication $P \Rightarrow Q$ and its converse $Q \Rightarrow P$ are not equivalent. Showing that is left as an exercise, along with proving the following negation rules for statements.

Theorem 3.2 (Negation rules for statements)**.** *For statements P and Q,*

$$\begin{aligned} \neg(P \wedge Q) &\iff \neg P \vee \neg Q, \\ \neg(P \vee Q) &\iff \neg P \wedge \neg Q, \\ \neg(P \Rightarrow Q) &\iff P \wedge \neg Q. \end{aligned}$$

3.3 Properties and quantifiers

Given a set X, a *property* on X is a sentence $P(x)$ that becomes either true or false (hence a statement) once a particular $x \in X$ is specified. For instance, if $X = \mathbb{Z}$ and $P(x)$ is "x is even," then $P(2)$ is true and $P(3)$ is false.

To turn a property into a statement, we use quantifiers. The *universal quantifier* $\forall$ means "for every" and the *existential quantifier* $\exists$ means "there exists." Continuing the example above, the statement

$$\forall x \in \mathbb{Z},\ P(x)$$

is false, since 3 is not even, while the statement

$$\exists x \in \mathbb{Z},\ P(x)$$

is true, since 2 is even.

Theorem 3.3 (Negation rules for quantifiers)**.** *For a property $P(x)$ on a set X,*

$$\neg(\forall x \in X,\ P(x)) \iff \exists x \in X,\ \neg P(x),$$
$$\neg(\exists x \in X,\ P(x)) \iff \forall x \in X,\ \neg P(x).$$

For example, $A \subseteq B$ means $\forall x\,(x \in A \Rightarrow x \in B)$. Its negation is

$$\exists x\,(x \in A \text{ and } x \notin B).$$

4 Proofs

4.1 What a proof is

A *theorem* is a statement that has been shown to be true, and a *proof* is the argument that shows it. Hammack's *Book of Proof* describes a proof as follows:

> "A proof of a theorem is a written verification that shows that the theorem is definitely and unequivocally true. A proof should be understandable and convincing to anyone who has the requisite background and knowledge."

A proof is typically written in ordinary prose plus mathematical notation. However, every step in a proof must be justified by something already accepted: a hypothesis of the theorem, a definition, a result established earlier, or a rule of logic. Proofs are often accompanied by pictures, but a picture usually does not replace a formal argument.

4.2 Direct proof

In a *direct proof* of an implication $P \Rightarrow Q$, we assume P and derive Q. In a direct proof of a universal statement $\forall x \in X,\ P(x)$, we let $x \in X$ be *arbitrary*, meaning that nothing is assumed about x beyond $x \in X$, and derive $P(x)$; since the argument used no property of x, it applies to every element of X.

Below is one example of a direct proof. It uses the following vocabulary: an integer n is *even* if $n = 2a$ for some $a \in \mathbb{Z}$, and *odd* if $n = 2a + 1$ for some $a \in \mathbb{Z}$. Every integer is even or odd, and no integer is both; we take that as given.

Lemma 4.1. *If n is an even integer, then $-5n - 3$ is an odd integer.*

Proof. Let n be an arbitrary even integer. Then $n = 2a$ for some $a \in \mathbb{Z}$, so

$$-5n - 3 = -5(2a) - 3 = -10a - 3 = 2(-5a - 2) + 1.$$

Now set $b = -5a - 2$. Since a is an integer, so is b, and we obtain $-5n - 3 = 2b + 1$. Hence $-5n - 3$ is odd. □

A *proof by contrapositive* of $P \Rightarrow Q$ is a direct proof of $\neg Q \Rightarrow \neg P$: we assume $\neg Q$ and derive $\neg P$. By Theorem 3.1 the two implications are equivalent, so proving either one proves the other. It is worth switching when the negations are the easier statements to work with.

Here is a second proof of Lemma 4.1, this time by contrapositive. Its hypothesis is that $-5n - 3$ is not odd, hence even, and its conclusion is that n is not even, hence odd.

Second proof of Lemma 4.1. We prove the contrapositive: if $-5n-3$ is an even integer, then n is an odd integer. Let n be an integer for which $-5n-3$ is even, so that $-5n-3 = 2b$ for some $b \in \mathbb{Z}$. Adding $6n+3$ to both sides gives

$$n = (-5n-3) + 6n + 3 = 2b + 6n + 3 = 2(b + 3n + 1) + 1.$$

Now set $c = b + 3n + 1$. Since b and n are integers, so is c, and the display reads $n = 2c + 1$. So n is odd, and therefore not even. By Theorem 3.1, this proves Lemma 4.1. □

4.3 Proof by contradiction

A proof by contrapositive requires the statement to be an implication. A *proof by contradiction* applies to a statement of any shape. To prove a statement S, we assume that S is false, and from that assumption we derive some statement C together with its negation $\neg C$. Whatever C may be, the statement $C \wedge \neg C$ is false, so the assumption cannot hold and S is true. The method is also called *reductio ad absurdum*: supposing the opposite of what we want reduces us to an absurdity.

The proof by contradiction is justified by the equivalence

$$S \iff \big(\neg S \Rightarrow (C \wedge \neg C)\big),$$

which holds for every statement C: showing that $\neg S$ leads to a contradiction is the same as showing that S is true. Verifying the equivalence with a truth table, as in the proof of Theorem 3.1, is left as an exercise.

When the statement to be proved is an implication $P \Rightarrow Q$, its negation is $P \wedge \neg Q$ by Theorem 3.2, so the proof begins by assuming both P and $\neg Q$. Which statement C will do the job is usually not known at the outset: we assume the negation, derive consequences, and stop once some statement and its negation are both on the page. Often the statement contradicted is one assumed at the start, a hypothesis or the negated conclusion, and sometimes it is a fact known independently.

Lemma 4.2. *Let $x \in \mathbb{R}$. If $x \leq \varepsilon$ for every real number $\varepsilon > 0$, then $x \leq 0$.*

Proof. The hypothesis P is that $x \leq \varepsilon$ for every $\varepsilon > 0$, and the conclusion Q is that $x \leq 0$. By Theorem 3.2, to assume the implication false is to assume P and $\neg Q$. Therefore, we start by assuming that

$$x \leq \varepsilon \ \text{ for every } \varepsilon > 0, \qquad \text{and} \qquad x > 0.$$

By the assumption on the right, $x/2 > 0$. The assumption on the left holds for every $\varepsilon > 0$, so it holds for $\varepsilon = x/2$, giving

$$x \leq \frac{x}{2}.$$

Subtracting $x/2$ from both sides gives $x/2 \leq 0$, and multiplying by 2 gives $x \leq 0$. The statement C is $x \leq 0$: we have just derived it, and its negation $x > 0$ is one of our assumptions. So C is both true and false, a contradiction. □

4.4 Proof by induction

Some statements come indexed by the natural numbers: for each $n \in \mathbb{N}$ there is a statement $P(n)$, and the claim is that $P(n)$ holds for every n. The following theorem reduces that claim, which involves infinitely many statements, to two checks.

Theorem 4.3 (Principle of mathematical induction)**.** *For each $n \in \mathbb{N}$, let $P(n)$ be a statement. Suppose that*

1. *$P(1)$ is true, and*
2. *for every $n \in \mathbb{N}$, $P(n) \Rightarrow P(n+1)$.*

Then $P(n)$ is true for every $n \in \mathbb{N}$.

Condition (1) is the *base step* and condition (2) is the *induction step*; the assumption $P(n)$ made while proving the induction step is the *induction hypothesis*. The two conditions act like a line of dominoes: the first one falls by (1), and each one that falls knocks over the next by (2), so all of them fall.

The induction step is itself a universal implication, so it is proved directly: let $n \in \mathbb{N}$ be arbitrary, assume $P(n)$, and derive $P(n+1)$.

Lemma 4.4 (Bernoulli's inequality)**.** *Let $x \in \mathbb{R}$ with $x > -1$. Then for every $n \in \mathbb{N}$,*

$$(1+x)^n \geq 1 + nx.$$

Proof. Fix $x > -1$ and let $P(n)$ be the displayed inequality.

Base step. For $n = 1$ both sides equal $1 + x$, so $P(1)$ holds.

Induction step. Let $n \in \mathbb{N}$ be arbitrary and assume $P(n)$, that is, $(1+x)^n \geq 1+nx$ (*induction hypothesis*). Multiply the induction hypothesis by $1 + x$, which is positive because $x > -1$, to obtain

$$(1+x)^{n+1} = (1+x)^n(1+x) \geq (1+nx)(1+x) = 1 + (n+1)x + nx^2.$$

Since $nx^2 \geq 0$, we have $(1+x)^{n+1} \geq 1 + (n+1)x$, which is $P(n+1)$. Thus $P(n) \Rightarrow P(n+1)$ for every $n \in \mathbb{N}$.

By Theorem 4.3, $P(n)$ holds for every $n \in \mathbb{N}$. $\square$